CLAIMS

1.   A method for categorizing text comprising the steps of:

dividing the text into sentences;

parsing the sentences into one or more noun phrases;

converting words in the noun phrases into networks of word relationships;

and

analyzing the word relationship networks to determine the influence of

each word.

2.   The method of claim 1 wherein the step of parsing the sentences into one or more noun phrases comprises the step of substituting disambiguated nouns in place of pronouns which are related to text analysis.

3.   The method of claim 1 wherein the step of parsing the sentences into one or more noun phrases further comprises the step of converting plural words to their singular form.

4.   The method of claim 1 wherein the step of converting the word or words into networks of relationships comprises the step of linking all sequentially occurring noun phrases within a sentence.

5.   The method of claim 4 wherein the step of converting the words into networks of word relationships further comprises the step of linking all possible pairs of words in the noun phrases having three or more words.

6.   The method of claim 1 wherein the step of analyzing the word relationship networks

to determine the influence of each word comprises the step of determining influence by utilizing between centrality.

7.   The method of claim 1 wherein the step of analyzing the word relationship networks to determine the influence of each word comprises the step of determining influence by utilizing the following formula:

$$ I = \frac{g_{jk}\ (i)/g_{jk}}{[(N-1)(N-2)/\ 2]} $$

where I is the influence of a word (i) in the text (T) where g_{jk} is the number of shortest paths connecting the jth and k^{th} words, $g_{jk}(i)$ is the number of those paths containing word (i), and N is the number of words in the network.

8.   A method for analyzing text comprising the steps of:

dividing the text into sentences;

parsing the sentences into one or more noun phrases;

converting one or more words within each of the noun phrases into networks of relationships between words;

analyzing the networks to determine the influence for each word; and

applying the analyzed networks to perform a specific analysis task.

9.   The method of claim 8 wherein the step of parsing the sentences into one or more noun phrases further comprises the step of substituting disambiguated nouns in place of pronouns which are relevant to text analysis.

10.   The method of claim 8 wherein the step of parsing the sentences into one or more

noun phrases further comprises the step of converting plural words to their singular form.

11. The method of claim 8 wherein the step of converting the word or words into networks of relationships comprises the step of linking all sequentially occurring noun phrases within a sentence.

12. The method of claim 11 wherein the step of converting the words into networks of relationships between words further comprises the step of linking all possible pairs of words in the noun phrases having three or more words.

13. The method of claim 8 wherein the step of analyzing the networks to determine the influence of each word comprises the step of determining influence by utilizing between centrality.

14. The method of claim 8 wherein the step of analyzing the networks to determine the influence of each word comprises the step of determining influence by utilizing the following formula:

$$I = \frac{g_{jk}(i)/g_{jk}}{[(N-1)(N-2)/2]}$$

where I is the influence of a word (i) in the text (T) where g_{jk} is the number of shortest paths connecting the jth and k^{th} words, $g_{jk}(i)$ is the number of those paths containing word (i), and N is the number of words in the network.

15. The method of claim 8 where the step of applying the analyzed network to perform a specific analysis task comprises the step of applying the analyzed network to

perform at least one of visualization of the network to understand text, spatial modeling of resonance scores, information retrieval, and thematic analysis of collections.

16. A method for determining resonance based on common words in two sets of text comprising the step of utilizing the following formula:

$$WR_{AB} = \sum_{i=1}^{N(A)} \sum_{j=1}^{N(B)} I_i^A \bullet I_j^B \bullet \alpha_{ij}^{AB}$$

where WR_{AB} is the word resonance between texts A and B, $\{w_1^A, w_2^A, \ldots w_{N(A)}^A)$ are unique words for text A after parsing into phrases in accordance with claim 8 where $N(A)$ is the number of unique words in text A, $\{I_1^A, I_2^A, \ldots I_{N(A)}^A\}$ are influence scores calculated in accordance with claim 14 for the unique words in text A, $\{w_1^B, w_2^B, \ldots w_{N(B)}^B)$ are unique words for text B after parsing into phrases in accordance with claim 8 where $N(B)$ is the number of unique words in text B, $\{I_1^B, I_2^B, \ldots I_{N(B)}^B\}$ are influence scores calculated in accordance with claim 14 for the unique words in text B. and indicator function α^{AB}_{ij} is equal to 1 if w_i^A and w_j^B are the same words, and the indicator function is equal to zero if w_i^A and w_j^B are not the same words.

17. The method of claim 16 further comprising the step of determining standardized resonance based on common words in texts A and B comprising the step of utilizing the following formula:

$$WR'_{AB} = WR_{AB} / \sqrt{\sum_{i=1}^{N(A)} (I_i^A)^2 \bullet \sum_{j=1}^{N(B)} (I_j^B)^2}$$

where WR$_{AB}$' is the standardized word resonance between texts A and B, WR$_{AB}$ is the actual word resonance between texts A and B, $\sum_{i=1}^{N(A)} \left(I_i^A \right)^2$ is the sum of all influence scores for the unique words in text A squared, and $\sum_{j=1}^{N(B)} \left(I_j^B \right)^2$ is the sum of all influence scores for the unique words in text B squared.

18.   A method for determining pair resonance based on common word pairs in two sets of text comprising the step of utilizing the following formula, where influence is calculated according to claim 14:

$$PR_{AB} = \sum_{i=1}^{N(A)-1} \left(\sum_{j=i+1}^{N(A)} \left(\sum_{k=1}^{N(B)-1} \left[\sum_{l=k+1}^{N(B)} P_{ij}^A \bullet P_{kl}^B \bullet \beta_{ijkl}^{AB} \right] \right) \right)$$

where PR$_{AB}$ is the word pair resonance between texts A and B, P$_{ij}^A$ is the frequency weighted pair influence of words i and j in text A and is equal to $I_i^A \bullet I_j^A \bullet F_{ij}^A$ where F$_{ij}^A$ is the number of times that w$_i^A$ and w$_j^A$ co-occur in text A, P$_{ij}^B$ is the frequency weighted pair influence of words k and l in text B and is equal to $I_k^B \bullet I_l^B \bullet F_{kl}^B$ where F$_{kl}^B$ is the number of times that w$_k^B$ and w$_l^B$ co-occur in text B, and indicator function β_{ijkl}^{AB} is equal to 1 if the two word sets (w$_i^A$, w$_j^A$) and (w$_k^B$, w$_l^B$) are equivalent and if F$_{ij}^A$ and F$_{kl}^B$ both are equal to one, otherwise the indicator is zero.

19.   The method of claim 18 further comprising the step of determining standardized resonance based on common word pairs in texts A and B comprising the step of utilizing the following formula:

$$PR'_{AB} = PR_{AB} \ / \ \sqrt{\left(\sum_{i=1}^{N(A)-1} \sum_{j=i+1}^{N(A)} \left(P_{ij}^A\right)^2 \right)} \cdot \sqrt{\left(\sum_{k=1}^{N(B)-1} \sum_{l=k+1}^{N(B)} \left(P_{kl}^B\right)^2 \right)}$$

where PR'$_{AB}$ is the standardized word pair resonance between texts A and B and PR$_{AB}$ is the actual word pair resonance between texts A and B.

20. A method for searching two or more texts utilizing resonance scores obtained in accordance with claim 16.

21. A method for searching two or more texts utilizing resonance scores obtained in accordance with claim 17.

22. A method for searching two or more texts utilizing resonance scores obtained in accordance with claim 18.

23. A method for searching two or more texts utilizing resonance scores obtained in accordance with claim 19.

24. A method for modeling two or more texts utilizing resonance scores obtained in accordance with claim 16.

25. A method for modeling two or more texts utilizing resonance scores obtained in accordance with claim 17.

26. A method for modeling two or more texts utilizing resonance scores obtained in accordance with claim 18.

27. A method for modeling two or more texts utilizing resonance scores obtained in accordance with claim 19.

28. A method for analyzing text comprising the steps of:

    a)   compartmentalizing the text into defined units;

    b)   categorizing the defined units by:

        parsing the units into one or more noun phrases each comprising one or more words;

        converting the word or words into networks of relationships between words;

        analyzing the networks of word associations to determine the structural influence of each word; and

    c)   applying the analyzed network to perform a specific analysis task.

29. The method of claim 28 wherein the step of compartmentalizing the text into defined units comprises the step of breaking down the text into sentences.

30. The method of claim 28 wherein the step of parsing the units into one or more noun phrases further comprises the step of substituting disambiguated nouns in place of pronouns which are relevant to text analysis.

31. The method of claim 28 wherein the step of parsing the units into one or more noun phrases further comprises the step of converting plural words to their singular form.

32. The method of claim 28 wherein the step of converting the word or words into networks of relationships comprises the step of linking all sequentially occurring noun phrases within a defined unit.

33. The method of claim 32 wherein the step of converting the word or words into

networks of relationships further comprises the step of linking all possible pairs of words in those noun phrases having three or more words.

34. The method of claim 28 wherein the step of analyzing the network to determine the structural influence of each word comprises the step of determining structural influence by utilizing betweeness centrality.

35. The method of claim 28 wherein the step of analyzing the network to determine the structural influence of each word comprises the step of determining structural influence by utilizing the following formula:

$$I = \frac{g_{jk}\ (i)/g_{jk}}{[(N-1)(N-2)/\ 2]}$$

where I is the influence of a word (i) in the text (T) where g_{jk} is the number of shortest paths connecting the jth and k^{th} words, $g_{jk}(i)$ is the number of those paths containing word (i), and N is the number of words in the network.

36. The method of claim 1 where the step of applying the analyzed network to perform a specific analysis task comprises the step of applying the analyzed network to perform at least one of visualization of the network to understand text, spatial modeling of resonance scores, information retrieval, and thematic analysis of collections.

37. A method for determining resonance based on common words in two sets of text comprising the step of utilizing the following formula:

$$WR_{AB} = \sum_{i=1}^{N(A)} \sum_{j=1}^{N(B)} I_i^A \bullet I_j^B \bullet \alpha_{ij}^{AB}$$

where WR_{AB} is the word resonance between texts A and B, $\{w_1^A, w_2^A, \ldots w_{N(A)}^A\}$ are unique words for text A after parsing into phrases in accordance with claim 28 where N(A) is the number of unique words in text A, $\{I_1^A, I_2^A, \ldots I_{N(A)}^A\}$ are influence scores calculated in accordance with claim 35 for the unique words in text A, $\{w_1^B, w_2^B, \ldots w_{N(B)}^B\}$ are unique words for text B after parsing into phrases in accordance with claim 28 where N(B) is the number of unique words in text B, $\{I_1^B, I_2^B, \ldots I_{N(B)}^B\}$ are influence scores calculated in accordance with claim 35 for the unique words in text B. and indicator function α_{ij}^{AB} is equal to 1 if w_i^A and w_j^B are the same words, and the indicator function is equal to zero if w_i^A and w_j^B are not the same words.

38. The method of claim 37 further comprising the step of determining standardized resonance based on common words in texts A and B comprising the step of utilizing the following formula:

$$WR'_{AB} = WR_{AB} / \sqrt{\sum_{i=1}^{N(A)}(I_i^A)^2 \bullet \sum_{j=1}^{N(B)}(I_j^B)^2}$$

where WR_{AB}' is the standardized word resonance between texts A and B, WR_{AB} is the actual word resonance between texts A and B, $\sum_{i=1}^{N(A)}(I_i^A)^2$ is the sum of all influence scores for the unique words in text A squared, and $\sum_{j=1}^{N(B)}(I_j^B)^2$ is the sum of all influence scores for the unique words in text B squared.

39. A method for determining pair resonance based on common word pairs in two sets of text comprising the step of utilizing the following formula, where influence is calculated in accordance with claim 35:

$$PR_{AB} = \sum_{i=1}^{N(A)-1} \left(\sum_{j=i+1}^{N(A)} \left(\sum_{k=1}^{N(B)-1} \left[\sum_{l=k+1}^{N(B)} P_{ij}^A \bullet P_{kl}^B \bullet \beta_{ijkl}^{AB} \right] \right) \right)$$

where PR_{AB} is the word pair resonance between texts A and B, P_{ij}^A is the frequency weighted pair influence of words i and j in text A and is equal to $I_i^A \bullet I_j^A \bullet F_{ij}^A$ where F_{ij}^A is the number of times that w_i^A and w_j^A co-occur in text A, P_{ij}^B is the frequency weighted pair influence of words k and l in text B and is equal to $I_k^B \bullet I_l^B \bullet F_{kl}^B$ where F_{kl}^B is the number of times that w_k^B and w_l^B co-occur in text B, and indicator function β_{ijkl}^{AB} is equal to 1 if the two word sets (w_i^A, w_j^A) and (w_k^B, w_l^B) are equivalent and if F_{ij}^A and F_{kl}^B both are equal to one, otherwise the indicator is zero.

40. The method of claim 39 further comprising the step of determining standardized resonance based on common word pairs in texts A and B comprising the step of utilizing the following formula:

$$PR'_{AB} = PR_{AB} \bigg/ \sqrt{\left(\sum_{i=1}^{N(A)-1} \sum_{j=i+1}^{N(A)} \left(P_{ij}^A\right)^2 \right)} \bullet \sqrt{\left(\sum_{k=1}^{N(B)-1} \sum_{l=k+1}^{N(B)} \left(P_{kl}^B\right)^2 \right)}$$

where PR'_{AB} is the standardized word pair resonance between texts A and B and PR_{AB} is the actual word pair resonance between texts A and B.

41.   A method for searching two or more texts utilizing resonance scores obtained in accordance with claim 37.

42.   A method for searching two or more texts utilizing resonance scores obtained in accordance with claim 38.

43.   A method for searching two or more texts utilizing resonance scores obtained in accordance with claim 39.

44.   A method for searching two or more texts utilizing resonance scores obtained in accordance with claim 40.

45.   A method for modeling two or more texts utilizing resonance scores obtained in accordance with claim 37.

46.   A method for modeling two or more texts utilizing resonance scores obtained in accordance with claim 38.

47.   A method for modeling two or more texts utilizing resonance scores obtained in accordance with claim 39.

48.   A method for modeling two or more texts utilizing resonance scores obtained in accordance with claim 40.